Alternative workflows for plant proteomic analysis†

Joohyun Lee and Bret Cooper*

DOI: 10.1039/b612326k

High-throughput separations are intrinsic to the detection and analysis of peptides and proteins by mass spectrometry (MS). Together, efficient separation and MS can lead to the identification of thousands of proteins in a sample, cell or tissue and help build proteome maps that can be used to define a cell type or cellular state. Although 2D gels have been successfully used to separate proteins for subsequent MS analysis, alternative separation efficiencies and, consequently deeper results could be obtained with HPLC or other separation techniques that improve throughput. This highlight is aimed toward plant scientists who have special separation needs due to the nature of plant cells and who could benefit from knowing options and requirements for adopting alternative separation protocols. Through the various sample processing and protein separation strategies, plant biologists should be able to improve the quality of their proteomic reference maps and gain new information about the proteins that define plant cells.

Introduction

If any of the proliferating neologistic biology terms appended with an -omics suffix are to be substantiated as scientific disciplines, then genomics and proteomics are the best candidates. Etymologically derived from chromosome and now scientifically realized as the order of base pairs within one, genomics has been legitimized by the entrenched technology platforms enabling high-throughput DNA sequencing and the wonderful biological insight provided by the related research. Minimal sets of genes that define life have been described[1] and genomics has led to unparalleled advancements in phylogenetics,[2–4] genetics[5,6] and breeding.[7] Of course, genomics spawned proteomics as a word and as a discipline that is empowered by high-throughput protein separation techniques and identification by mass spectrometry (MS). By itself, genomics provides a list of genes in a cell but this list contributes little to the knowledge of which genes are at play and what their functions are. However, coupled with proteomics, an understanding of protein accumulation and organization can emerge.

For the many organisms whose genomes have been sequenced, the assembly

USDA-ARS, Soybean Genomics and Improvement Laboratory, 10300 Baltimore Ave., Bldg 006 Rm 213, Beltsville, MD 20705 E-mail: cooperb@ba.ars.usda.gov

† Electronic supplementary information (ESI) available: Peptide sequences, Mascot scores, protein groups and their probabilities for *A. thaliana* leaves analyzed by MudPIT. See DOI: 10.1039/b612326k

Joohyun Lee

Dr Joohyun Lee is a Postdoctoral Research Associate in the Soybean Genomics and Improvement Laboratory, USDA-ARS, Beltsville, MD. His current research is focused on shotgun proteomics approaches used to detect and identify plant proteins involved in responses to plant pathogens. He majored in agronomy and received an MS in Plant Genetics and Breeding from the Seoul National University in the Republic of Korea and obtained a PhD from Louisiana State University.

Bret Cooper

Dr Bret Cooper is a Research Plant Pathologist in the Soybean Genomics and Improvement Laboratory, USDA-ARS, Beltsville, MD. His current research includes discovering disease-causing proteins from infectious fungi, the defense-related proteins from the host plants and the molecular interactions between them. Scientists within his laboratory commonly use mass spectrometry to study these proteins as well as identify pathogens in infected plants. Previously, Dr Cooper worked at the Torrey Mesa Research Institute, La Jolla, CA where he was involved in several rice and Arabidopsis thaliana functional genomics programs. He majored in Biology at Washington University in St. Louis, received a PhD in Plant Pathology from the University of California, Riverside and performed post-doctoral studies at the Scripps Research Institute, La Jolla, CA.

of protein accumulation reference maps that provide a global snapshot of the protein constituents of cells in various states and conditions has begun.[8–11] Plants are no exception.[12,13] Plant biology has greatly benefited from genomics but the rate of proteomics discovery in plants has not kept up with genome sequencing, partly because of difficulties in obtaining proteins from the many developmentally different yet spatially convergent cell types in various plant tissues. Plant cells, compared to animal cells, are also compressed by vacuoles and rigid cell walls traversed by plasmodesmata and are crowded by many membranous plastids. These differences introduce new challenges for protein extraction and separation not entirely encountered by researchers who pioneered proteomics on yeast, bacteria and cultured animal and human cells. Thus, we offer this overview of MS workflow options to plant scientists who are ever more interested in defining plant proteomes but are unaware of which separation techniques will meet their needs. We will also show that these considerations can make a difference in the amounts and types of plant proteins that can be resolved. For this discussion, we primarily focus on the model plants *Arabidopsis thaliana* and rice, whose completed genome reference sequences are essential to the proteomic analysis of these and other plants.[14,15]

2D-PAGE

The 2-dimensional polyacrylamide gel electrophoresis (2D-PAGE) technique for the separation of complex protein mixtures was developed in the 1970s,[16] and remains a dominant method for resolving proteins. In this method proteins are separated by their isoelectric point through a pH-gradient gel matrix[17] (IPG) and then by molecular weight in the second dimension. The extra dimension of separation and the added area of the gel matrix allows for the clear separation of several thousand proteins. 2D-PAGE attained its full potential when researchers used mass spectrometry to identify the separated proteins, also known as spots, within the gel.[18] Excision of a protein spot can be performed by hand with a razor blade, and the following procedures for destaining, in-gel protein digestion with a protease such as trypsin and the elution of peptides are cookbook procedures that can be easily performed.[19] The peptide mixtures can then be sent to a service lab to obtain affordable and reliable results (Fig. 1A). Usually, the labs will use matrix-assisted laser-desorption ionization (MALDI)-based time-of-flight (TOF) MS to determine peptide mass fingerprints (PMF) of the peptides from the spot.[20,21] Alternatively, electrospray ionization-based tandem MS (MS/MS) is used.[22,23] In both cases, PMFs and tandem mass spectra are usually searched against protein sequence database references and the result produced indicates a match between an observed spectrum and a virtual spectrum produced from the database.[24,25] Based upon the quality of the match, the identity of the protein can be deduced.

By subjecting treated and control samples or samples from different time points to 2D-PAGE, a differential comparison of protein accumulation can also be obtained.[26] Quantitation can be added to the mix by using a variety of different *in vitro* labeling techniques (ICAT; isotope-coded affinity tag),[27] or using sensitive and differential gel stains (DIGE; difference gel electrophoresis).[28,29] All of these steps, with the exception of running 2D gels, are amenable to various types of automation which allows hundreds of differentially resolved proteins to be quickly extracted, processed and identified.[19]

There is a respectable amount of proteomics research reporting the protein accumulation, profiling patterns and identity of proteins from various *A. thaliana*,[30–32] and rice[33–35] plant tissues. Depending on the tissue, treatment

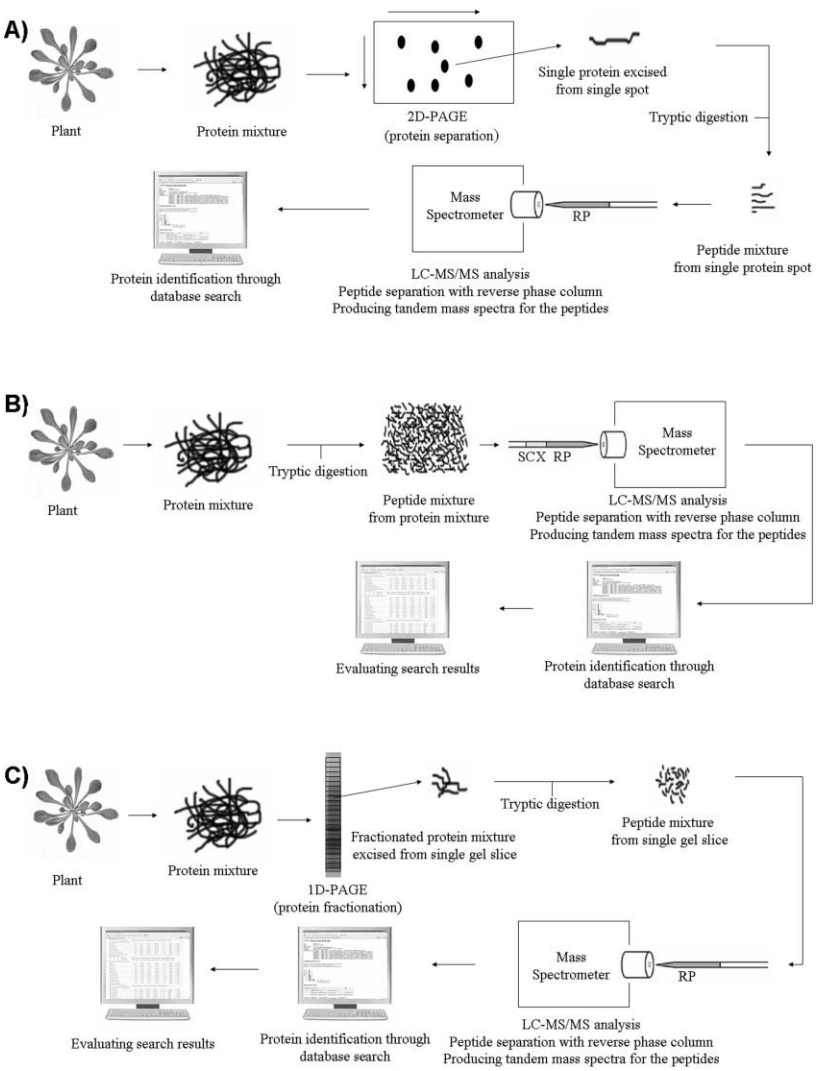

Fig. 1 Various proteomics workflows. (A) 2D-PAGE; (B) MudPIT; (C)1D-Gel-LC-MS/MS.

or plant species, hundreds to a few thousand proteins have been resolved and in general, many of these studies serve as catalogues for the accumulating proteins in plant cells or organelles.[36–41] There are several examples whereby Japanese and Australian groups have independently constructed 2D-PAGE databases that serve as proteomic maps. The results are reported as digitized images of the 2D gels whereby scientists around the world can study any protein spot (http://gene64.dna.affrc.go.jp/RPD/main_en.html; http://semele.anu.edu.au/2d/2d.html). In the Australian rice anther protein database, there are more than 4000 protein spots within the pI range 4–11 and molecular weight range of 6–122 kDa, corresponding to ~10% of the rice genome.[42] The Japanese database provides 21 reference maps of proteins from specific developmental tissues consisting of more than 13 000 spots.[43] Complementing whole tissue studies are analyses of subcellular fractions specifically containing plasma membranes, chloroplasts, mitochondria or nuclei.[36–41] More than anything, subcellular fractionation helps decrease the complexity of the whole tissue protein extracts.[37]

2D-PAGE is widely used by plant scientists for protein separation because the equipment and materials for this technique are readily available and easily applied. 2D-PAGE allows scientists to proceed quickly with protein discovery and whet their proteomics appetite. Those with a greater hunger can layer in differential analysis, automation or quantitation to delve deeper into the proteome. Eventually, however, the researchers reach the resolution limitations of 2D-PAGE. Usually, 1000 to 2000 proteins can be displayed. However, with plants, the highly abundant ribulose bisphosphate carboxylase/oxygenase complex (RuBisCO) proteins, which parallel blood plasma albumin in terms of protein concentration, overwhelm many of the low abundance proteins that would otherwise be clearly resolved.[44,45] Unfortunately, no good RuBisCO protein depletion strategies exist as they do for serum albumin. To overcome RuBisCO abundance, researchers have used IPG strips with overlapping pI ranges to reduce sample complexity.[46,47] This increases the number of gels that need to be run and consequently increases labor, cost, and time. The lack of automation for running 2D gels, the bottleneck of image analysis and spot picking and the fact that the physical integrity of the gels deteriorates after several hundred spots are extracted create formidable and steadfast barriers to a high-throughput workflow. Even when researchers are willing to roll up their sleeves to plow through a 2D gel workload, the lack of resolving power for basic, hydrophobic and membrane spanning proteins could be the weakest point to 2D-PAGE.[48] These limitations restrict the detection of nearly 30% of all cellular proteins, and will certainly impact proteomics research on plant cells packed with various and specialized membranous structures such as thylakoids and plasmodesmata.[49]

Multidimensional protein identification technology (MudPIT)

A clear advancement to the proteomics workflow is MudPIT, which refers to the unique HPLC protein separation method that allows the high-throughput analysis of proteins by MS.[50] Basic, hydrophobic and membrane-spanning proteins have a better chance of being separated by HPLC, as long as they can be obtained for analysis. The initial benefits of this workflow were described by Washburn *et al*, (2001)[51] who discovered more membrane spanning proteins with MudPIT compared to 2D-PAGE.[52,53]

Unlike 2D gels where selected proteins are digested after separation, MudPIT analysis requires that all proteins in a sample are digested into peptides before the separation step (Fig. 1B). The separated peptides are then sequentially eluted into the mass spectrometer and analyzed.[51] This method enables high-throughput analysis and a user can walk away from the mass spectrometer while the peptides are eluted. MudPIT is also compatible with quantitative and differential comparative analysis, which is made possible by labeling proteins *in vivo*[54,55] or by labeling peptides *in vitro*.[56–58] However, one drawback is that multiple analyses are usually required to circumvent random sampling errors associated with all MudPIT-style assays.[59]

While MudPIT opens up the range for detecting proteins, the technique introduces a new set of problems that have to be reconciled. First, since a sample is loaded directly into the mass spectrometer, detergents most commonly used to isolate hydrophobic proteins must be avoided since these are readily ionized and can mask results of the less-easily ionized peptides.[60] MS-compatible detergents are now available and users have to remember to incorporate these into their protein extraction protocols.[61] Second, data output comprises tens of thousands of tandem mass spectra, which necessitates the use of software such as Sequest and Mascot,[24,25] or *de novo* sequencing programs[62–64] to deduce the amino acid sequence information from the spectra. Realistically, a cluster of computer processors running the algorithms is required to process the vast amounts of spectra in a reasonable amount of time. Next, since the original mixture of proteins was digested into peptides, the protein information must be regenerated. Reassembly of peptides into proteins can be quite daunting in light of the fact that many of these proteins can share the same sequences. Without knowing which proteins appeared in the starting sample, the only reasonable way to generate a candidate list is to obey rules of parsimony such that the smallest, most-logical non-redundant set of proteins is assembled from the peptides.[65,66] Freely available programs such as DTASelect[67] and DBParser[68] organize the peptides into sets of candidate proteins and ProteinProphet[69] and PANORAMICS[70] perform a similar organization with the added benefit of incorporating a probability model that can be used to ascertain the likelihood that a protein assembly is correct.

The MudPIT workflow follows a basic blueprint, and there are variations thereof. In the original setup, peptides are separated on columns coupled to an HPLC pump with the orifice of the column directed toward the source of the ion trap of the mass spectrometer.[71] Columns are usually constructed from fused-silica capillaries pulled to a fine tip capable of spraying liquid drops of ionized peptides into the mass spectrometer source.[72] The column is typically packed with strong cation exchange (SCX) resin and reverse phase (RP) resin,

peptides are loaded onto the SCX resin and the column is placed in-line with the HPLC pump. By alternating the flow of salt and organic solvents the peptides are eluted from the column, first by charge and then by polarity.[50] Multiple separation steps can be introduced simply by increasing the concentration of the salt in the solvent at each step, which selectively elutes classes of peptides mobilized under those concentrations. Each salt step is then followed by the same increasing gradient of organic solvent which further separates those peptides. Variations of this pattern such as designing RP-SCX-RP columns and consecutive SCX-RP configurations can aid sample cleanup or give extended separations.[73] Other variations in this workflow include collecting off-line fractions of the peptide mixture as they elute from an SCX column, and then separating each SCX fraction by RP.[74] This variation often utilizes an autosampler and is as nearly as amenable to high-throughput and hands-free separation as the conventional MudPIT architecture.

MudPIT typically requires many custom-made materials, including columns, column packing pressure cells, HPLC tube fittings and platforms that configure the column to the HPLC pump. The specialization, the bioinformatics platforms and expertise required keep this technology out of many plant researchers' hands. As a result, there are only a few reports on the high-throughput analysis of proteins from plants. The most notable is by Koller *et al.* (2002), who identified more than 2 300 proteins from the leaves, roots, and seeds of rice by performing only a handful of experiments.[13] Such results are very tantalizing and are prompting more researchers to venture into using the MudPIT workflow for proteomic research.

1D-Gel-LC-MS/MS

An emerging method gaining popularity brings the advantages of MudPIT-style separations to researchers who can not easily acquire the customized resources needed for MudPIT. Those more capable of running gels may be intrigued by a method that combines 1D gel separations with RP liquid chromatography (Fig. 1C). Here proteins are first separated by size on standard polyacrylamide gels,[75,76] or by isoelectric point on IPG strips normally used for the first dimensional separation in 2D-PAGE.[77] The 1D-gel separation step takes the place of the SCX separation of MudPIT and results in reducing the complexity of a sample. After separation, the lane of the gel or the strip containing the proteins is extracted and then divided into 32 slices or some other number compatible with multiple sample handling procedures. The gel slice is treated similarly to spots excised from 2D gels and the peptides are separated on an integrated and reusable RP column coupled to a standard HPLC pump. The RP eluent is then analyzed by MS/MS. Because the gel slices contain many different proteins which are subsequently digested into peptides, protein reassembly is also required. Nevertheless, the main benefit to this workflow is that common 1D gel resources are available to a broader cross-section of researchers and the results obtained by separating proteins on 1D gels are parallel to MudPIT.[75] The downside to this technique is that the offline protein fractionation and gel handling utilizes more labor and time than MudPIT. But researchers willing to sacrifice the elbow grease can survey many proteins in one experiment. Plant researchers have already successfully used this method for subcellular, organelle and membrane proteomics research.[38,78–80] We have also discovered that many of the detergents that are not compatible with MudPIT—detergents commonly used to extract proteins from unique plant membrane structures—can be sufficiently removed by 1D gel separation and subsequent clean-up methods.

Comparing workflows for protein discovery

In this penultimate section, we present an example of the successes that can be achieved when performing a proteomic analysis of *A. thaliana* leaves using MudPIT.[81] Briefly, whole leaves were pulverized in liquid nitrogen and the material was washed in trichloracetic acid/acetone and then dried to produce a crude protein lyophilate. Proteins were resolubilized in Tris buffer, urea, and dodecyl-beta-maltoside (DDM), a detergent that can be easily removed before MS/MS.[82,83] The solubilized proteins were digested with trypsin. Peptides were eluted from a SCX-RP capillary column using a 12-step separation method and detected by a Thermo LCQ Deca XP ion trap mass spectrometer. The tandem mass spectra were analyzed with Mascot and compared to proteins in v. 6.0 of the *A. thaliana* protein sequence database obtained from www.arabidopsis.org. The subsequent data were evaluated by the probability-based protein assembly software PANORAMICS and all proteins exceeding a 95% confidence level for identification were accepted. By combining the results from two replicate experiments, 594 non-redundant proteins were identified. As many as 11% of those proteins appeared to have membrane spanning domains when searched against ARAMEMNON, a plant membrane protein database (http://aramemnon.botanik.uni-koeln.de/index.ep).†

To dig deeper into the cell, we prepared subcellular fractions of the leaves to reduce sample complexity.[84] Fresh leaves were pulverized in liquid nitrogen and Tris buffer, and then centrifuged at 1000 × g. The resulting pellet (CW) containing nuclei, cell walls and chloroplasts was washed three times in 2% DDM and boiled in urea. The 1000 × g supernatant was then centrifuged at 30 000 × g to produce an organelle pellet (P30) and a supernatant (S30). The P30 pellet was also washed in DDM and boiled in urea. Proteins from the S30 fraction were precipitated with acetone. Proteins from all fractions were analyzed by the same MudPIT workflow as before. Two replicates for each were performed and the data from each replicate combined. We found 501, 481 and 701 non-redundant proteins in the respective CW, P30 and S30 fractions.† Together, this resulted in 1204 non-redundant proteins. Among them 123, 116, 44 membrane-spanning proteins were identified from the respective CW, P30 and S30 fractions resulting in 198 (16%) non-redundant membrane-spanning proteins. Combining the results from the crude protein analysis and the analysis of the subcellular fractionations, we detected 1369 leaf proteins at a 95% confidence level.† Of this whole set, only 448 proteins (33%) were identified by a single peptide. The PANORAMICS probability-model allows single peptide identifications but has several requirements, such as the peptide must be

assigned very high Mascot scores and the peptide sequence must be a unique (distinct) identifier for the protein. The remaining proteins were identified by more than one peptide whose multiplicity increased the probability that a protein was identified.

With respect to time, it took 2 days to prepare the samples, 1 day to prepare custom columns and 8 days to run the samples. The bioinformatics analysis required 1 day. In total, we needed 12 work days to discover 1369 leaf proteins from 8 samples. We do not believe that we could have achieved the same results in the same amount of time using 2D-PAGE. Assuming that we can extract 500 protein spots from a 2D gel using a robot, we estimate from prior experience that it would take 5 days to run the gel, excise spots, digest and elute peptides, 11 days to perform LC-MS/MS, and 2 days to analyze data. Multiply this by four for the different sample preparations and any required replicates and it becomes quite apparent that MudPIT can resolve as many or more proteins in a much shorter amount of time. Finally, we would not have expected to have been able to detect as many membrane-spanning proteins by 2D-PAGE.

As an aside, we want to state that these results do not imply that MudPIT is a superior workflow. Instead, we show these results to illustrate some advantages to adopting an alternative workflow. We staunchly believe that 2D-PAGE, MudPIT, and 1D-Gel separations are often complementary and necessary for full proteomic characterization. In the future, we will describe the benefits of employing all three methods for plant proteomic analysis.

Concluding remarks

The dynamics of the proteome are dependent on the development stage of the cell and the environmental conditions in which it exists.[85] For rice, which is estimated to harbor 35 000–60 000 genes,[6,14] it is possible that 500 000 different forms of proteins could be produced from them. The dynamic range of these proteins could also be quite excessive; the rare proteins are present at a level of 10 molecules per cell whereas the most abundant proteins are present at levels of 10^7.[85] It is therefore seemingly impossible to ascertain how many proteins are present within any specific cell type, detect how they are modified, determine their concentrations, or measure their turnover rates. In the past, protein accumulation and dynamics have only been measured accurately for one or a few proteins at a time and this is usually with the aid of antibodies or specific conjugation to fluorescent molecules.[86] It is still not yet possible to evaluate a proteome with the same type of resolution. Indeed, to complete a plant proteome map is to reach an indefinable goal. However, MS-based proteomics allows us to approach that goal and that is a good start for an emerging -omics field.

Acknowledgements

The authors wish to thank Drs Neerav Padliya and Mark Tucker at USDA-ARS, Beltsville, MD for their help in preparing this Highlight article.

References

1 A. Goffeau, B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin and S. G. Oliver, *Science*, 1996, **274**, 546.
2 F. C. Chen and W. H. Li, *Am. J. Hum. Genet.*, 2001, **68**, 444.
3 L. L. Lutfiyya, N. Xu, R. L. D'Ordine, J. A. Morrell, P. W. Miller and S. M. Duff, *J. Plant Physiol.*, 2006, DOI: 10.1016/j.jplph.2006.04.014.
4 M. Premzl, J. E. Gready, L. S. Jermiin, T. Simonic and J. A. Marshall Graves, *Mol. Biol. Evol.*, 2004, **21**, 2210.
5 P. H. Dear, A. T. Bankier and M. B. Piper, *Genomics*, 1998, **48**, 232.
6 International Rice Genome Sequencing Project, *Nature*, 2005, **436**, p. 793.
7 S. Hittalmani, A. Parco, T. V. Mew, R. S. Zeigler and N. Huang, *Theoret. Appl. Genet.*, 2000, **100**, 1121.
8 J. Alonso, J. M. Rodriguez, L. A. Baena-Lopez and J. F. Santaren, *J. Proteome Res.*, 2005, **4**, 1636.
9 N. Hall, M. Karras, J. D. Raine, J. M. Carlton, T. W. Kooij, M. Berriman, L. Florens, C. S. Janssen, A. Pain, G. K. Christophides, K. James, K. Rutherford, B. Harris, D. Harris, C. Churcher, M. A. Quail, D. Ormond, J. Doggett, H. E. Trueman, J. Mendoza, S. L. Bidwell, M. A. Rajandream, D. J. Carucci, J. R. Yates, 3rd, F. C. Kafatos, C. J. Janse, B. Barrell, C. M. Turner, A. P. Waters and R. E. Sinden, *Science*, 2005, **307**, 82.
10 A. Madi, S. Mikkat, B. Ringel, M. Ulbrich, H. J. Thiesen and M. O. Glocker, *Proteomics*, 2003, **3**, 1526.
11 D. Nedelkov, U. A. Kiernan, E. E. Niederkofler, K. A. Tubbs and R. W. Nelson, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 10852.
12 R. Holmes-Davis, C. K. Tanaka, W. H. Vensel, W. J. Hurkman and S. McCormick, *Proteomics*, 2005, **5**, 4864.
13 A. Koller, M. P. Washburn, B. M. Lange, N. L. Andon, C. Deciu, P. A. Haynes, L. Hays, D. Schieltz, R. Ulaszek, J. Wei, D. Wolters and J. R. Yates, 3rd, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 11969.
14 S. A. Goff, D. Ricke, T. H. Lan, G. Presting, R. Wang, M. Dunn, J. Glazebrook, A. Sessions, P. Oeller, H. Varma, D. Hadley, D. Hutchison, C. Martin, F. Katagiri, B. M. Lange, T. Moughamer, Y. Xia, P. Budworth, J. Zhong, T. Miguel, U. Paszkowski, S. Zhang, M. Colbert, W. L. Sun, L. Chen, B. Cooper, S. Park, T. C. Wood, L. Mao, P. Quail, R. Wing, R. Dean, Y. Yu, A. Zharkikh, R. Shen, S. Sahasrabudhe, A. Thomas, R. Cannings, A. Gutin, D. Pruss, J. Reid, S. Tavtigian, J. Mitchell, G. Eldredge, T. Scholl, R. M. Miller, S. Bhatnagar, N. Adey, T. Rubano, N. Tusneem, R. Robinson, J. Feldhaus, T. Macalma, A. Oliphant and S. Briggs, *Science*, 2002, **296**, 92.
15 The Arabidopsis Genome Initiative, *Nature*, 2000, **408**, p. 796.
16 P. H. O'Farrell, *J. Biol. Chem.*, 1975, **250**, 4007.
17 B. Bjellqvist, K. Ek, P. G. Righetti, E. Gianazza, A. Gorg, R. Westermeier and W. Postel, *J. Biochem. Biophys. Methods*, 1982, **6**, 317.
18 J. R. Yates, 3rd, *J. Mass Spectrom.*, 1998, **33**, 1.
19 K. Gevaert, T. Houthaeve and J. Vandekerckhove, *EXS*, 2000, **88**, 29.
20 M. Karas and F. Hillenkamp, *Anal. Chem.*, 1988, **60**, 2299.
21 B. Valot, S. Gianinazzi and D. G. Eliane, *Phytochemistry*, 2004, **65**, 1721.
22 J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong and C. M. Whitehouse, *Science*, 1989, **246**, 64.
23 P. A. Haynes, N. Fripp and R. Aebersold, *Electrophoresis*, 1998, **19**, 939.
24 J. Eng, A. L. McCormack and J. R. Yates, 3rd, *J. Am. Soc. Mass Spectrom.*, 1994, **5**, 976.
25 D. N. Perkins, D. J. Pappin, D. M. Creasy and J. S. Cottrell, *Electrophoresis*, 1999, **20**, 3551.
26 S. Cui, F. Huang, J. Wang, X. Ma, Y. Cheng and J. Liu, *Proteomics*, 2005, **5**, 3162.
27 F. Schmidt, B. Dahlmann, K. Janek, A. Kloss, M. Wacker, R. Ackermann, B. Thiede and P. R. Jungblut, *Proteomics*, 2006, **6**, 4622.
28 S. Amme, A. Matros, B. Schlesier and H. P. Mock, *J. Exp. Bot.*, 2006, **57**, 1537.
29 S. Komatsu, X. Zang and N. Tanaka, *J. Proteome Res.*, 2006, **5**, 270.
30 K. Gallardo, C. Job, S. P. Groot, M. Puype, H. Demol, J. Vandekerckhove and D. Job, *Plant Physiol.*, 2001, **126**, 835.
31 M. Kamo, T. Kawakami, N. Miyatake and A. Tsugita, *Electrophoresis*, 1995, **16**, 423.

32 A. Tsugita, M. Kamo, T. Kawakami and Y. Ohki, *Electrophoresis*, 1996, **17**, 855.
33 S. Komatsu, A. Muhammad and R. Rakwal, *Electrophoresis*, 1999, **20**, 630.
34 R. Rakwal and S. Komatsu, *Electrophoresis*, 2000, **21**, 2492.
35 A. Tsugita, T. Kawakami, Y. Uchiyama, M. Kamo, N. Miyatake and Y. Nozu, *Electrophoresis*, 1994, **15**, 708.
36 M. S. Bae, E. J. Cho, E. Y. Choi and O. K. Park, *Plant J.*, 2003, **36**, 652.
37 V. Kruft, H. Eubel, L. Jansch, W. Werhahn and H. P. Braun, *Plant Physiol.*, 2001, **127**, 1694.
38 A. H. Millar and J. L. Heazlewood, *Plant Physiol.*, 2003, **131**, 443.
39 J. B. Peltier, G. Friso, D. E. Kalume, P. Roepstorff, F. Nilsson, I. Adamska and K. J. van Wijk, *Plant Cell*, 2000, **12**, 319.
40 T. A. Prime, D. J. Sherrier, P. Mahon, L. C. Packman and P. Dupree, *Electrophoresis*, 2000, **21**, 3488.
41 D. Rouquie, J. B. Peltier, M. Marquis-Mansion, C. Tournaire, P. Doumas and M. Rossignol, *Electrophoresis*, 1997, **18**, 654.
42 N. Imin, T. Kerim, J. J. Weinman and B. G. Rolfe, *Proteomics*, 2001, **1**, 1149.
43 S. Komatsu, K. Kojima, K. Suzuki, K. Ozaki and K. Higo, *Nucleic Acids Res.*, 2004, **32**, D388.
44 C. C. des Francs, H. Thiellement and D. de Vienne, *Plant Physiol.*, 1985, **78**, 178.
45 S. T. Kim, K. S. Cho, Y. S. Jang and K. Y. Kang, *Electrophoresis*, 2001, **22**, 2103.
46 S. Hoving, H. Voshol and J. van Oostrum, *Electrophoresis*, 2000, **21**, 2617.
47 R. Wildgruber, A. Harder, C. Obermaier, G. Boguth, W. Weiss, S. J. Fey, P. M. Larsen and A. Gorg, *Electrophoresis*, 2000, **21**, 2610.
48 V. Santoni, M. Molloy and T. Rabilloud, *Electrophoresis*, 2000, **21**, 1054.
49 R. Douce and J. Joyard, *Annu. Rev. Cell Biol.*, 1990, **6**, 173.
50 D. A. Wolters, M. P. Washburn and J. R. Yates, 3rd, *Anal. Chem.*, 2001, **73**, 5683.
51 M. P. Washburn, D. Wolters and J. R. Yates, 3rd, *Nat. Biotechnol.*, 2001, **19**, 242.
52 B. Futcher, G. I. Latter, P. Monardo, C. S. McLaughlin and J. I. Garrels, *Mol. Cell. Biol.*, 1999, **19**, 7357.
53 M. Perrot, F. Sagliocco, T. Mini, C. Monribot, U. Schneider, A. Shevchenko, M. Mann, P. Jeno and H. Boucherie, *Electrophoresis*, 1999, **20**, 2280.
54 L. M. de Godoy, J. V. Olsen, G. A. de Souza, G. Li, P. Mortensen and M. Mann, *Genome Biol.*, 2005, **7**, R50.
55 N. Gustavsson, B. Greber, T. Kreitler, H. Himmelbauer, H. Lehrach and J. Gobom, *Proteomics*, 2005, **5**, 3563.
56 R. Chen, S. Pan, E. C. Yi, S. Donohoe, M. P. Bronner, J. D. Potter, D. R. Goodlett, R. Aebersold and T. A. Brentnall, *Proteomics*, 2006.
57 J. Hu, J. Qian, O. Borisov, S. Pan, Y. Li, T. Liu, L. Deng, K. Wannemacher, M. Kurnellas, C. Patterson, S. Elkabes and H. Li, *Proteomics*, 2006.
58 X. Chen, S. W. Cushman, L. K. Pannell and S. Hess, *J. Proteome Res.*, 2005, **4**, 570.
59 H. Liu, R. G. Sadygov and J. R. Yates, 3rd, *Anal. Chem.*, 2004, **76**, 4193.
60 D. Drexler, D. J. Barlow, P. Falk, J. Cantone, D. Hernandez, A. Ranasinghe, M. Sanders, B. Warrack and F. McPhee, *Anal. Bioanal. Chem.*, 2006, **384**, 1145.
61 M. Cadene and B. T. Chait, *Anal. Chem.*, 2000, **72**, 5655.
62 A. Shevchenko, S. Sunyaev, A. Loboda, A. Shevchenko, P. Bork, W. Ens and K. G. Standing, *Anal. Chem.*, 2001, **73**, 1917.
63 A. Shevchenko, I. Chernushevic, A. Shevchenko, M. Wilm and M. Mann, *Mol. Biotechnol.*, 2002, **20**, 107.
64 A. Shevchenko, I. Chernushevich, M. Wilm and M. Mann, *Methods Mol. Biol.*, 2000, **146**, 1.
65 A. I. Nesvizhskii and R. Aebersold, *Mol. Cell. Proteomics*, 2005, **4**, 1419.
66 S. Carr, R. Aebersold, M. Baldwin, A. Burlingame, K. Clauser and A. Nesvizhskii, *Mol. Cell. Proteomics*, 2004, **3**, 531.
67 D. L. Tabb, W. H. McDonald and J. R. Yates, 3rd, *J. Proteome Res.*, 2002, **1**, 21.
68 X. Yang, V. Dondeti, R. Dezube, D. M. Maynard, L. Y. Geer, J. Epstein, X. Chen, S. P. Markey and J. A. Kowalak, *J. Proteome Res.*, 2004, **3**, 1002.
69 A. I. Nesvizhskii, A. Keller, E. Kolker and R. Aebersold, *Anal. Chem.*, 2003, **75**, 4646.
70 J. Feng, D. Q. Naiman and B. Cooper, *Mol. Cell. Proteomics,* submitted.
71 A. J. Link, J. Eng, D. M. Schieltz, E. Carmack, G. J. Mize, D. R. Morris, B. M. Garvik and J. R. Yates, 3rd, *Nat. Biotechnol.*, 1999, **17**, 676.
72 C. L. Gatlin, G. R. Kleemann, L. G. Hays, A. J. Link and J. R. Yates, 3rd, *Anal. Biochem.*, 1998, **263**, 93.
73 H. W. McDonald, R. Ohi, D. T. Miyamoto, T. J. Mitchison and J. R. Yates, 3rd, *Int. J. Mass Spectrom.*, 2002, **219**, 245.
74 S. Wienkoop, M. Glinski, N. Tanaka, V. Tolstikov, O. Fiehn and W. Weckwerth, *Rapid Commun. Mass Spectrom.*, 2004, **18**, 643.
75 L. Breci, E. Hattrup, M. Keeler, J. Letarte, R. Johnson and P. A. Haynes, *Proteomics*, 2005, **5**, 2018.
76 U. K. Laemmli, *Nature*, 1970, **227**, 680.
77 B. J. Cargile, J. L. Bundy and J. L. Stephenson, Jr., *J. Proteome Res.*, 2004, **3**, 1082.
78 M. Ferro, D. Salvi, H. Riviere-Rolland, T. Vermat, D. Seigneurin-Berny, D. Grunwald, J. Garin, J. Joyard and N. Rolland, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 11487.
79 D. J. Maltman, W. J. Simon, C. H. Wheeler, M. J. Dunn, R. Wait and A. R. Slabas, *Electrophoresis*, 2002, **23**, 626.
80 A. Marmagne, M. A. Rouet, M. Ferro, N. Rolland, C. Alcon, J. Joyard, J. Garin, H. Barbier-Brygoo and G. Ephritikhine, *Mol. Cell. Proteomics*, 2004, **3**, 675.
81 B. Cooper, W. M. Garrett and K. B. Campbell, *Proteomics*, 2006, **6**, 2477.
82 L. Florens and M. P. Washburn, *Methods Mol. Biol.*, 2006, **328**, 159.
83 S. Luche, V. Santoni and T. Rabilloud, *Proteomics*, 2003, **3**, 249.
84 B. Cooper and J. A. Dodds, *J. Gen. Virol.*, 1995, **76**(12), 3217.
85 H. Barbier-Brygoo and J. Joyard, *Plant Physiol. Biochem.*, 2004, **42**, 913.
86 C. T. Dooley, T. M. Dore, G. T. Hanson, W. C. Jackson, S. J. Remington and R. Y. Tsien, *J. Biol. Chem.*, 2004, **279**, 22284.

